Utilistic reduction in sociology

Kuipers, Theo A.F.

Theo A.F. Kuipers

UTILISTIC REDUCTION IN SOCIOLOGY: THE CASE OF COLLECTIVE GOODS

1. INTRODUCTION

One aim of methodological individualism in sociology is to explain social regularities on the basis of some regularity in the behaviour of individuals. An additional aim may be to explain such an individual regularity in turn on the basis of the assumption that the individuals intend to maximize expected utility, or at least that they behave as if they intend to do so. Explanations of social regularities exemplifying both features will here be called utilistic reduction.

In this paper we shall present a detailed account of one such utilistic reduction, viz. the explanation of Mancur Olson's hypothesis 'the larger the group, the farther it will fall short of providing an optimal ammount of a collective good' (Olson (1965), p.35)[1].

Olson's own explanation, indicated in the title of his book 'The Logic of Collective Action', has already been streamlined and simplified by S. Lindenberg (1982). However, Lindenberg's account still has many shortcomings, as we have pointed out extensively in our (1983c). Compared with the expositions of Olson and Lindenberg, our account in (1983c) brings to the fore a number of hidden elements and necessary refinements, and it is shown that the resulting explanation has exactly the structure of so-called heterogeneous reduction.

In this paper we will report the main findings of our (1983c); but in several respects the analysis will also be improved.

Section 2 spells out the three fundamental steps of the reduction of Olson's hypothesis.

In Section 3 we will discuss three additional issues, of increasing generality. First, we will show that the expectations of the individuals with respect to the transformation process of collective behaviour into social effects are largely correct, and hence rational, in the light of the objective transformation rules. Nevertheless - and this is well-known in the 'Olson-literature'- the utilistic be-

W. Balzer et al. (eds.), Reduction in Science, 239–267.

haviour of the individuals based on these expectations leads to rather irrational consequences. Second, the treated utilistic reduction is essentially an as-if explanation. We will argue that a new explication of intentional explanations, presented in our (1983b), provides the adequate means to transform any as-if-utilistic reduction into a genuine intentional-utilistic reduction. Third, and finally, in any type of reduction in sociology (i.e. utilistic or not) transformation rules will be necessary. We will discuss a number of types of such rules, and possibilities to explain them.

In the remainder of this section we shall clarify and illustrate the notion of heterogeneous reduction.

When we talk in this paper about reduction we mean throughout deductive micro-reduction, as opposed to something like approximative reduction or to forms of reduction not involving different ontological levels. In the light of these restrictions a theory (or law, or set of laws) T' is said to be reduced to T if and only if 1) T is of a 'lower' or 'more fundamental' ontological level than T' and 2) T' is derivable from T, possibly with the aid of transformation rules relating some concepts of the two theories. We speak of heterogeneous reduction if (non-analytic) transformation rules are necessary to accomplish the deduction, otherwise we speak of homogeneous reduction [2].

Homogeneous reduction turns out to be decomposable into two steps. In the first step, the individual step, some individual law is derived on the basis of the reducing theory and special assumptions. An individual law is here understood as any uniform feature in the behaviour of the individuals of the reducing theory. In the second homogeneous step, the aggregation step, this individual law is aggregated in some way or other, for instance, with or without using statistical assumptions. This aggregation leads to some aggregated law, which completes a homogeneous reduction.

In the case of heterogeneous reduction we also start with these two homogeneous steps. In an additional, third step, the transformation step, transformation rules are applied to the aggregated law, in order to derive the law intended to be reduced.

We will illustrate heterogeneous reduction, and hence also homogeneous reduction, by a global sketch of the kinetic reduction of the ideal gas law, i.e. the reduction of the ideal gas law (PV = RT) to the kinetic theory of gases.

Individual step. From the kinetic gas theory, in particular Newton's laws of motion, and the assumption that the molecules collide elastically with the wall, one derives the individual law that the momentum-exchange m_e in a collision between a molecule and the wall is given by:

$$m_e = 2mv_w$$

where m is the mass of the molecule and v_w the velocity of the molecule in the direction perpendicular to the wall.

Aggregation step[3]. Through aggregation, using some statistical assumptions, we can derive from this individual law the 'perfect gas law' (for one mole of gas):

PGL $PV = (2/3)N\bar{u}$

Here V indicates the volume of the gas, P its pressure, $\bar{u}$ the mean kinetic energy of the molecules and N Avogadro's number of molecules in a mole. With some qualifications, which are irrelevant here, we can say that PGL uses only 'aggregated' quantities and hence is an aggregated law.

Transformation step. The usual form of the transformation rule is:

$$\bar{u} = (3/2)kT = (3/2)(R/N)T$$

where T indicates the (absolute) temperature of the gas, k Boltzmann's constant (= R/N) and R the ideal gas constant. It is easy to check that with this rule it is possible to derive from PGL the ideal gas law:

IGL $PV = RT$

In Kuipers (1982) we have given a detailed analysis of the transformation step [4], showing that 1) the realistic restriction to 'asymptotic' behaviour can be made, 2) the apparent symmetry (PGL can also be derived from IGL and the transformation rule) can be removed, 3) an additional transformation rule is required for pressure, and, finally, 4) both transformation rules can be formulated in such a way that they express so-called ontological identities.

2. THE REDUCTION OF OLSON'S HYPOTHESIS

2.1. Preparations

Let us now turn to Olson's hypothesis. First we need to make a number of terminological remarks. Point of departure are *groups* of people defined in some way. Somebody belonging to a group is called a *member* of that group. A good is called a *collective good* for a group if all members of the group will benefit from it, whether or not one participates in the pursuit of that good. One very common kind of participation is to be a (contribution paying) member of the organization (e.g. a labour union) pursuing the collective good. However, we will speak of *participants* without any specific form of participation in the pursuit of the collective good in mind. The number of participants divided by the total number of members of the group (the size of the group) will be called the *degree of participation*.

If the collective good is not a matter of degree but an 'all or nothing' object (e.g. a bridge), the crucial notion will be the *chance* that the collective good will be realized, the *chance of realization* for short. On the other hand, if the collective good is something that can be realized in principle at any level, the crucial notion will be the *level* at which the collective good is realized, the *level of realization*. In the first case we will speak of the *chance-variant*, in the second case of the *level-variant*. Mixed variants are of course also possible. Due to the unclear notion of 'chance' in the present context, the chance-variant is more vague than the level-variant. Nevertheless, we will present the analysis in the chance-variant.

Now we are in a position to reformulate Olson's hypothesis as follows:

OH The larger the group the smaller the chance of realization.

The basic idea of the whole reduction is the decomposition of OH into the following two hypotheses:

OH-I The larger the group the lower the degree of participation.

OH-II The lower the degree of participation the smaller the chance of realization.

In the homogeneous part of the reduction, OH-I is then derived as an aggregated regularity on the basis of an individual regularity. This individual regularity says roughly that an individual stops participating if 'his' group reaches a certain size, and this is explained as the consequence of maximizing expected utility.

In the heterogeneous step OH is derived from OH-I with OH-II as a transformation rule, transforming 'degree of participation' into 'chance of realization'.
Although the aggregation and transformation step may sound rather trivial in this sketch, it will turn out that all three steps require careful treatment in order to get a rigorous deductive account.

2.2. Individual step

The individual regularity to be derived in the first step will be explained by a number of hypotheses about the considerations individuals might use in choosing between participation and non-participation. Although the presentation will suggest that the individuals are consciously deliberating, this is not essential. That is, our aim and claim in this section is only an 'as-if explanation'. As announced in the introduction, we will discuss how to transform such an as-if-explanation in Section 3.

Since any set of individuals could constitute, in principle, a real group, having or not (yet) having a collective good, we will call every set of individuals a group. In many cases, however, it will be clear that the following implicit assumption is made:
'suppose that this set constitutes a real group'.

To begin with, the following abbreviations should be interpreted in this light:

X	a (fixed type of) collective good
I	the universe of all individuals (under consideration)
i	a variable for individuals, i.e. $i \in I$
G	a variable for sets of individuals, groups, i.e. $G \subseteq I$
P(i,G)	i participates in the pursuit of X for G
N(i,G)	not-P(i,G)
P(G)	$\{i \in G \ / \ P(i,G)\}$, the subset of participants of G
N(G)	$\{i \in G \ / \ N(i,G)\} = G-P(G)$
n(G)=n	the size of G
m(G)=m	the size of P(G), $o \leq m \leq n$
m/n	the degree of participation in G
EU-A(i)	i's (subjective) expected utility of (i's) action A

As a general hypothesis we will assume the classical principle of utility theory: individuals maximize their expected utility in the choice between participation and non-participation (the P/N-choice, for short):

MH <u>Maximization hypothesis</u>
For all G and all i in G
P(i,G) iff EU-P(i,G) $\geq$ EU-N(i,G)

The equation in MH will be called the <u>participation equation</u>. Note that MH contains the harmless but convenient assumption that individuals participate when their expected utilities of the two actions are equal.

To be able to apply MH we have of course to introduce <u>specification hypotheses</u> for the utilities and expectations of the individuals. We start with the following

<u>Utility hypotheses</u>

U1 For all G and all i in G
i's utility of X for G is positive and independent of G; it is represented by $U_X(i)$.

U2 For all G and all i in G
i's subjective costs, i.e. negative utility, of participation in the pursuit of X for G are positive and independent of G; it is represented by $C_P(i)$.

Of course we assume that no other utility than those specified in U1/2 (i.e. U1 and U2) are involved. Note that it is not difficult to think of situations where one or more aspects of U1/2 do not seem very reasonable. The most delicate assumption is surely contained in U2, viz. the fact that the subjective costs of participation are assumed to be independent of the group and hence of the size of the group. It may well be, for example, that the (financial) contribution decreases with increasing size, and this is likely to decrease the subjective costs as well.

Turning to the expectations of the individuals we will first formulate the relevant hypotheses for an arbitrary individual i. These hypotheses will also be called the <u>subjective transformation rules</u>. Lateron we will assume, in addition, <u>objective transformation rules</u>, i.e. rules factually governing the transformation process of collective behaviour into social phenomena. We start with:

E1 $R_i = R_i(n,m)$
i.e. i's estimate R_i of the chance of realization of X for G depends only on the size of G and P(G), i.e. n and m, and so we can denote it by $R_i(n,m)$.

Let $b_i(n,m)$ denote $R_i(n,m+1) - R_i(n,m)$, $m < n$. For obvious reasons, $b_i(n,m)$ will be called the <u>marginal effect</u> according to i. Our next assumption is:

E* $b_i(n,m) = b_i(n)$
i.e. $b_i(n,m)$ does not depend on m, and hence can be represented as $b_i(n)$.

For convenience we introduce the conjunction

E*1 E1 and E*

The two remaining assumptions are dependent on E*:

E*2 $b_i(n) > o$
i.e. the marginal effect, according to i, is always positive

E*3 $b_i(n) > b_i(n+1)$
i.e. the marginal effect decreases with increasing size of the group according to i

Of these assumptions, E*2 is quite plausible and E*3 will turn out to be crucial.

We combine and generalize these assumptions for all individuals:

E*H <u>Expectation hypothesis</u>
For all i, R_i satisfies E*1/2/3, i.e. E*1, E*2 and E*3.

In order to formulate and prove the intended individual regularity we have to define a technical notion.

<u>Def. 1</u> i has <u>switch (group) size</u> S(i) if
for all G, for which i in G, holds
P(i,G) iff $n(G) < S(i)$
(and hence N(i,G) iff $n(G) \geq S(i)$)

Now it is possible to prove the following theorem:

<u>Th.1</u> MH, U1/2, E*H imply together:
SL <u>Switch Law</u>: all i have a (specific) switch group size, i.e. for all i there is a natural number S(i) such that S(i) is the switch group size of i.

Proof An individual i in G is supposed to argue as follows according to MH, U1/2 and E1: if there are z participants, besides myself, then the utilities I can expect are respectively

$EU\text{-}P(i,G) = R_i(n,z+1).\ U_X(i) - C_P(i),$

$EU\text{-}N(i,G) = R_i(n,z).\ U_X(i)$

The resulting participation equation reduces, on the basis of E*, to

$b_i(n) \geq C_P(i)/U_X(i),$

in which the hypothetical number z has disappeared [5]. According to U1/2 and E*2 both sides of this equation are positive. Moreover, the quotient on the right side is positive. But, according to E*3, $b_i(n)$ decreases with increasing n. Hence, there will exist, as a rule, a (finite) switch group size, viz. the smallest number n for which $b_i(n) < C_P(i)/U_X(i)$. For the case that $b_i(n)$ does not 'pass' the quotient, despite E*3, we define of course $S(i) = \infty$.

Q.e.d.

It is interesting to note that the notion of the switch size is a purely behavioural concept, i.e. it does not refer to utilities and expectations. Consequently, the switch law SL is a purely behavioural regularity, i.e. a regularity in individual behaviour. In what follows it will also become clear that the role of the utilistic hypotheses (MH, U1/2, E*H) in the utilistic reduction of Olson's hypothesis is restricted to the presented explanation of SL. These facts have two interesting consequences. 1) If a rival explanation of SL can be given, on the basis of different utilistic hypotheses or on the basis of non-utilistic hypotheses, the rest of the story could remain the same. 2) As we have said before, we only claim that the preceding story gives at least an as-if- explanation. To transform the whole resulting as-if-utilistic reduction into a genuine intentional-utilistic reduction it will be sufficient to transform the given as-if explanation of SL into an intentional explanation of SL.

2.3. Aggregation step

Although OH-I may already sound plausible in the light of the switch law, an exact derivation of it is still problematic. For, consider the following more explicit formulation of OH-I (leaving out here and in what follows the obvious quantifiers):

OH-I If G_2 is larger than $G_1 (n_2 > n_1)$ then the degree of participation in G_2 is lower than that in $G_1 (m_2/n_2 < m_1/n_1)$.

Now, SL does not allow the derivation of OH-I, for, on the one hand, OH-I is, as it stands, a claim about any two groups of different size, whereas, on the other hand, SL does not imply any restriction on the switch sizes of different individuals.

Consequently, in order to derive 'something like OH-I' from SL, the derivation should deal, in some way or another, with the <u>same</u> individuals. If we succeed in this, this 'something like OH-I' may well be called an <u>explication</u> of OH-I, for it will be a revision of OH-I which captures, in a precise way, what we can defensibly mean by, and what is nevertheless 'close to', the initial formulation.

For these purposes we introduce the notions of <u>fusion</u> and <u>division</u> of groups, again hinting at, but formally not requiring, real groups.

<u>Def. 2</u> G_3 is a <u>fusion</u> of G_1 and G_2 $(G_3 = G_1 \oplus G_2)$ if $G_1 \cap G_2 = \emptyset$ and $G_1 \cup G_2 = G_3$. G_1 and G_2 are then said to constitute a <u>division</u> of G_3.

Now, if G_3 is a fusion of G_1 and G_2, we can argue, on the basis of the switch law, as follows. There may be individuals who were participants in G_1 or G_2, for whom the larger size of G_3 reaches (and probably passes) their switch size. Hence, these individuals become non-participants in G_3. However, the converse cannot occur, i.e. SL excludes that there may have been non-participants in G_1 or G_2 who become participants in G_3.

In sum, it is possible to prove

<u>Th.2</u> SL implies:
<u>LF</u> <u>Law of Fusion</u>: if $G_3 = G_1 \oplus G_2$ then $P(G_3) \subseteq P(G_1) \cup P(G_2)$

One might claim that LF itself is already an explication of OH-I. But note that LF is a purely qualitative statement, whereas OH-I is quantitative, though only comparative. It turns out that a (mathematical) consequence of LF, hence also a consequence of SL, is in this respect closer to OH-I, viz.

Oh-Ie If $G_3 = G_1 \oplus G_2$ then $m_3 \leq m_1 + m_2$

That this is 'something like OH-I' becomes more clear from an equivalent 'mean-value-formulation' (the equivalence being due to the analytical fact that $n_3 = n_1 + n_2$ if $G_3 = G_1 \oplus G_2$), viz.

$$\text{If } G_3 = G_1 \oplus G_2 \text{ then } \frac{m_3}{n_3} \leq \frac{n_1}{n_1+n_2} \cdot \frac{m_1}{n_1} + \frac{n_2}{n_1+n_2} \cdot \frac{m_2}{n_2}$$

which reads: the degree of participation in a fused group is not higher than the weighted mean value of the degrees of participation in the fusing groups. Of course, the weakening of 'lower than' to 'not higher than' is the price of soundness.

Although the mean-value-formulation 'sounds' more like OH-I than the 'number-formulation', the former is just a clarifying alternative to the latter. However, for the explication of OH instead, we will be <u>forced</u> into such a mean-value-formulation (in terms of chances of realization), because there an equivalent number-formulation will not be available.

It will of course depend on the nature of groups and of the collective good whether the idea of fusion and division is purely theoretical or has some practical applications. If a group is defined as the set of all those individuals in a society who share interest in the collective good under consideration, the composition of this group only changes through 'natural' mutations. Hence, in this case of <u>unique</u> groups, fusion and division are purely hypothetical. If, however, groups are defined in another way, i.e. independent of the collective good, as for instance families, or, more general, communities and if the collective good is only related to each group seperately, then actual fusion and division becomes possible. Hence, for such <u>non-unique</u> groups experimental testing of the law of fusion and hence of OH-Ie is possible.

Given the structure of LF and OH-Ie it is reasonable to call them regularities in collective behaviour, or <u>collective regularities</u> for short. Given the way in which they could be derived from SL it is also reasonable to call them aggregated regularities. Moreover, it is also clear that the joint derivation of OH-Ie in the individual and the aggregation step satisfies all requirements for homogeneous reduction as described in the introduction.

2.4. <u>Transformation step</u>

The last step concerns the transformation of collective behaviour into social phenomena. In the individual step we made assumptions about the expectations individuals had concerning this transformation. In contrast to the there postulated subjective transformation rules, E*1/2/3, our concern here are objective transformation rules.

Let us first restate the verbal version of the main transformation rule:

OH-II The lower the degree of participation the smaller the chance of realization.

The following rule seems to be more or less implicit in OH-II and corresponds to E.1:

TR1 $R = R(n,m)$
i.e. R, the objective chance of realization of X for G, depends only on the size of G and P(G), i.e. n and m, and hence can be represented as R(n,m).

The objective marginal effect of participation b(n,m) is of course defined as $R(n,m+1) - R(n,m)$, $m < n$. Although we will not assume that the marginal effect depends only on n, we will assume that it is always positive (compare E*2):

TR2 $b(n,m) > o$

It is easy to check that TR2 is equivalent to

TR'2 if $o \leq m < m' \leq n$ then $R(n,m) < R(n,m')$

and this is rather close to a formal version of OH-II, assuming TR1. Hence, let us take TR1 and TR2 as a provisional explication of OH-II and let us see how far we can come, on the basis of them and OH-Ie, in deriving OH:

OH The larger the group the smaller the chance of realization.

We refer here to the explication OH-Ie of OH-I in fusion terms, which we had to design in order to be able to derive it from the switch law, for we want to retain of course this partial explanation in the inclusive explanation of OH.

To achieve this, it is clear that we will have to explicate OH in fusion terms as well. In the case of OH-I there were two equivalent formulations, a number- and a mean-value-formulation. However, since the chance of realization is not a matter of natural numbers, we are now forced into a mean-value-formulation:

OHe If $G_3 = G_1 \oplus G_2$ then

$$R(n_3,m_3) \leq \frac{n_1}{n_1+n_2} \cdot R(n_1,m_1) + \frac{n_2}{n_1+n_2} \cdot R(n_2,m_2)$$

Note that TR1 is used in the transition from OH to OHe.

It is easy to see that OH-Ie and TR1/2 do not yet imply OHe, but only:

If $G_3 = G_1 \oplus G_2$ then $R(n_3,m_3) \leq R(n_1+n_2,m_1+m_2)$

To derive OHe from this we need an additional transformation rule, of which the formulation in fusion terms reads:

TR3 $R(n_1+n_2,m_1+m_2) \leq \frac{n_1}{n_1+n_2} \cdot R(n_1,m_1) + \frac{n_2}{n_1+n_2} \cdot R(n_2,m_2)$

Putting

OH-IIe = TR1 & TR2 & TR3 = TR1/2/3

we have indicated the proof of

Th.3 OH-Ie and OH-IIe imply OHe

It has not been noted by Olson and Lindenberg that something like TR3 is a hidden assumption in the claim that (something like) OH can be explained.

It is important to realize that the need for TR3 has nothing to do with the fusion-formulations of OH-I and OH: the fusion-formulations of the latter two forces only a fusion-version of a straightforward idea, viz. that the chance of realization may not increase with increasing size of the group, when the degree of participation remains constant (despite OH-I!), or formally:

If $n < n'$ and $m/n = m'/n'$ then $R(n',m') \leq R(n,m)$.

In our (1983c), Section 3.5, we have illustrated the relevance of TR3 by presenting a possible counterexample to it, viz. in terms of fixed costs for the realization of a collective good for a group, i.e. costs which are independent of the size of the group. If there are such fixed costs, they need to be payed twice for two groups, but only once for a fused group. This creates the possibility that the chances for the collective good increase with fusion despite a possible decline of the degree of participation.

Although we have already hinted a little at the relation between the (objective) transformation rules TR1/2/3 and the subjective transformation rules E*1/2/3, the precise relation will be studied in Section 3.1.

In Section 3.3, where we will discuss types of (objective) transformation rules in general, we will also briefly comment on the nature of TR1/2/3. But it will already be clear by now that they are non-analytic, and hence that the transformation step is a truly heterogeneous step, from which it follows that the total explanation of OH (i.e. OHe) is indeed an example of heterogeneous (utilistic) reduction, as claimed in the introduction.

2.5. Survey

The deductive pattern of the treated example can be represented by

$$\text{TS}\ \frac{\text{AS}\ \dfrac{\text{IS}\ \dfrac{\text{MH} \ + \ \text{U1/2} \ + \ \text{E*H}}{\text{SL}}}{\text{OH-Ie}} \qquad \text{OH-IIe (=TR1/2/3)}}{\text{OHe}}$$

where IS stands for individual step, AS for aggregation step and TS for transformation step.

Before we lay down the general structure of utilistic reduction we will first state a division of regularities (or laws, or rules) which we have already suggested from time to time. Of course the formulations should be interpreted with careful flexibility.

Individual regularities or, more precisely, regularities in individual behaviour (such as the switch law): those and those circumstances lead to that and that individual behaviour.
Collective regularities or, more precisely, regularities in collective behaviour (such as OH-I): those and those circumstances lead to that and that collective behaviour.
Transformation rules (such as OH-II): that and that collective behaviour leads to that and that social phenomenon.
Social regularities (such as OH): those and those circum-

stances lead to that and that social phenomenon.

In these general terms we get as a general pattern:

UTILISTIC REDUCTION OF A SOCIAL REGULARITY		
IS	maximization hypothesis	specification hypotheses
AS	individual regularity	
TS	collective regularity	transformation rules
	social regularity	

If all transformation rules used in a utilistic reduction are analytic it is a case of homogeneous reduction. As soon as there is at least one empirical transformation rule it is a heterogeneous reduction.

3. RATIONALITY, INTENTIONAL INTERPRETATION AND TYPES OF TRANSFORMATION RULES

In this section we will deal with three topics of increasing generality. First, we will evaluate the rationality of the individualsin the case of the utilistic reduction of Olson's hypothesis concerning collective goods. Second we will show how any as-if-utilistic reduction can be transformed into an intentional-utilistic reduction. Finally, we will give a short survey of possible transformation rules for any type of reduction in sociology.

3.1. Rational expectations and irrational consequences

We will first investigate the relation between the objective and the subjective transformation rules, i.e. between TR1/2/3 of Section 2.3 and E*1/2/3 of Section 2.1. As we have seen, both rules play a crucial role in the reduction. If the postulated beliefs about the transformation process of the individuals, were incompatible with, or highly different from, the objective transformation rules, the reduction story would only have peculiar applications. But, if the beliefs of the individuals are largely correct in the light of the objective transformation rules, i.e. if their expectations are largely rational, there can be many natural situations where all reduction conditions are, roughly, satisfied.

Let us first note that the two sets of rules are compatible. That is, there is at least one (formal) example for the objective chance of realization R which satisfies TR1/2/3 as well as E*1/2/3. It is easy to verify that $R(n,m) = m/n$ is such an example.

Our main question, however, is the following: how much is true of E*1/2/3 if TR1/2/3 are true? In the light of the fact that TR1 trivially implies E1, this question reduces to: how much is true of E*, E*2 and E*3 if $R(n,m)$ satisfies TR2 and TR3?

Let us first formulate weaker versions of E*2 and E*3, viz.

E2 $b_i(n,m) > o$

E3 $$\frac{S_i(n,n)}{n} > \frac{S_i(n+1,n+1)}{n+1}$$

where $S_i(n,m)$ is defined as $R_i(n,m) - R_i(n,o)$, i.e. $\sum_{z=o}^{m-1} b_i(n,z)$.

It is easily checked that E*2 and E*3 are in fact the conjuntion of E2 and E3, respectively, with E* (i.e. $b_i(n,m) = b_i(n)$), in the same way as E*1 is, by definition, the conjunction of E1 and E*.

It is clear that TR1/2/3 do not imply E*. On the other hand, it is not only a trivial fact that they do imply E1, but also that they imply E2. Hence, the remaining question is: do TR1/2/3 imply E3? This is an important question, for E*3 is, as we have seen, a crucial element in the reduction, and it is clear that E3 grasps the main point of E*3 in this respect, viz. the average marginal influence decreases with increasing size of the group. The answer to this question is 'yes' and hence we have in total the following theorem.

Th.4 TR1/2/3 imply E1/2/3

Proof E1 and E2 are trivially implied, as already indicated. The general validity of TR3 obviously requires, substituting $R(n,m) = R(n,o) + S(n,m)$,

$$S(n_1+n_2, m_1+m_2) \leq \frac{n_1}{n_1+n_2} \cdot S(n_1,m_1) + \frac{n_2}{n_1+n_2} \cdot S(n_2,m_2) \tag{1}$$

and a similar condition for $R(n,o)$.
It is not difficult to check that

$$\frac{n'}{m'} S(n',m') \leq \frac{n}{m} S(n,m) \text{ for } n' \geq n,\ m' \geq m \tag{2}$$

is not only a sufficient but also a necessary condition for (1).
Substitution of $m' = n' = n+1$ and $m = n$ in (2) leads to

(3) $S(n+1,n+1) \leq S(n,n)$

From TR2 trivially follows that

(4) $S(n,n) > o$

Finally, (3) and (4) trivially imply E3 (for S). Q.e.d.

We summarize Th.4 provisionally in the slogan: the expectations of the individuals are largely rational. This point has not been made by Olson and Lindenberg. As to the remaining 'possible distance' between TR1/2/3 and E*1/2/3 the following points are relevant. If R satisfies, in addition to TR1/2/3, the independence condition corresponding to E*, i.e.

TR* $b(n,m) = b(n)$

it follows from Th.4 that R has all the properties E*1/2/3. Nevertheless, there need not be quantitative correspondence between $R_i(n,m)$ and $R(n,m)$ in this case.
However this may be, if TR* holds in addition to TR1/2/3, the individuals base their behaviour on qualitatively correct insights in the transformation process.

If TR* is approximately true the behaviour of the individuals is still based on more or less correct insights. If, however, TR* is fundamentally false (whereas TR1/2/3 are true) the assumption E* is no longer rational. Inspection of the proof of Th.1 shows that the decision to participate should then no longer be made independent of what the others are going to do. Hence, the above mentioned slogan, i.e. 'the expectations are largely rational', should be qualified with 'provided the objective marginal effect roughly depends only on the size of the group'. To be sure, the qualification will be true in many cases.

Despite the rational expectations of the individuals, their behaviour leads to rather irrational consequences. This irrational aspect of Olson's logic of collective action has been discussed extensively in the literature. It is also known under the heading 'the tragedy of the commons'. Taylor (1976) shows that the individuals are subject to (a generalized form of) the so-called prisoner's dilemma, with non-participation as the dominant strategy.

Here we will only show these irrational consequences for an extreme case. Let us first assume, in general, that the individuals have perfect knowledge about the transformation process concerning X, i.e.

PK For all i, $R_i(n,m) = R(n,m)$

Note that PK presupposes TR1 as well as E1. Let us further assume,

TR* $b(n,m) = b(n)$

as well as TR2 and TR3. From PK and Th.4 it follows that E*1/2/3 are all satisfied. Let us also assume U1/2.

Consider now a group G the size n of which ia larger than, or equal to, the switch size S(i) for all i in G. If they all behave according to MH, i.e. maximize their expected utility, nobody will participate, i.e. $P(G) = \emptyset$.

Inspection of the proof of Th.1 shows that all these assumptions imply:

(5) For all i in G, $b(n) < C_P(i) / U_X(i)$

Every individual i in G can now objectively expect the utility:

(6) $R(n,o)\ U_X(i)$

If everybody would, despite (5), participate in the pursuit of X, i.e. not maximize expected utility, but 'behave socially', each individual could then objectively expect the utility:

(7) $R(n,n)\ U_X(i) - C_P(i)$

This expected utility is, for all i in G, larger than that according to (6) if, using $nb(n) = R(n,n) - R(n,o)$,

(8) For all i in G, $nb(n)\ U_X(i) > C_P(i)$

It is important to note that the conjunction of (5) and (8), i.e.

(9) For all i in G, $b(n)\ U_X(i) < C_P(i) < nb(n)\ U_X(i)$

is a real possibility, despite its extreme character.

In the case of (9), the conclusion is straightforward: 'social behaviour' by everybody leads to a larger expected utility for everybody than 'maximizing expected utility behaviour' by everybody. It is intuitively clear that this paradoxical phenomenon will not only occur in the extreme case (9), but that it will also occur in some partial sense in less extreme cases.

From the fact that the foregoing story was based on PK it follows already that the irrational consequences of maximizing expected utility behaviour cannot be due to the lack of rationality of the expectations of the individuals. In the first half of this subsection we have seen that the expectations may be called rational, even without perfect knowledge, i.e. without quantitative correspondence. Also in these cases it will be possible to derive similar irrational consequences. Hence, under rather general conditions the source of the described kind of irrationality cannot be the expectations. It is of course also absurd to assume that the source could be the utility assignments U1/2, for they function as primitives in the diagnosis of irrationality. Hence, the remaining candidate, viz. the principle of maximizing expected utility itself, is indeed the source of irrationality.

3.2. Intentional interpretation

In this section we will deal with the intentional interpretation of an as-if-utilistic reduction. To be precise, we will show how the actions of the individuals can be explained intentionally in terms of the conscious pursuit of maximal utility or, for short, how they can be explained as intentional-utilistic.

Of course we do not claim that such an intentional interpretation is always realistic. The only thing which matters is that such an interpretation will be realistic in some cases and our problem is what we precisely mean by that in such cases. We will present our analysis in terms of the reduction of Olson's hypothesis, but it will be clear that the content of the example does not play a substantial role.

In the preceding paragraphs we have alluded to an intuitive notion of intentional explanation of actions which needs further explication. Now one might expect that this explication will be of a nomological nature, given the 'nomological approach' of reduction in this paper. In particular, one might expect that the current nomological explication of intentional explanations (NI-explanations) will provide the adequate intentional interpretation.

However, in the literature many different objections have been made to NI-explanations. We will present one such objection in some detail. Suppose P(i,G), i.e. i participates in the pursuit of X for G. For the whole of this section we will implicitly assume that i is a member of G, if i and G occur as free variables. For completeness' sake we will first specify the nomological *as-if* explanation, with the maximization hypothesis MH in the role of nomological premise.

$$\frac{\text{EU - P(i,G)} \geq \text{EU - N(i,G)} \quad \text{MH}}{\text{P(i,G)}}$$

This is an *as if* explanation because i's (conscious) wish of maximal utility does not play a role.

In the NI-explanation of P(i,G) the following *desiderative* premise is explicitly stated

WMU(i,G): i wishes (wished) maximal utility in the P/N-choice

Moreover, the *epistemic* premise

DU-P(i,G) EU-P(i,G) $\geq$ EU-N(i,G)

is interpreted as the (conscious) belief of i that his expected utility of participation 'dominates' that of non-participation. For convenience, not-DU-P(i,G) will be equated with DU-N(i,G). Finally, both premises are taken into account in the *new* nomological premise:

CPMU For all G and all i in G,
if WMU(i,G) then
P(i,G) iff DU-P(i,G)

CPMU states roughly that everybody is always consistent (C) in his pursuit (P) of maximum utility (MU). The resulting NI-explanation is clearly a deductive argument

$$\frac{\text{WMU(i,G)} \quad \text{DU-P(i,G)} \quad \text{CPMU}}{\text{P(i,G)}}$$

The main objection to this NI-explanation concerns the role of the nomological premise CPMU. It belongs to the nomological view that CPMU should be interpreted as an empirical (i.e. falsifiable) premise [6] and that the explanation is only adequate if there are good reasons to assume that all required premises are true, hence also CPMU. Suppose now that there have been observed one or more convincing counterexamples to CPMU, i.e. 'inconsistent individuals'. Hence, CPMU is false. Now it follows that P(i,G) cannot be explained intentionally, according to the NI-explanation, in terms of the pursuit of maximal utility, even if the particular individual i has nothing to do with these counterexamples.

The obvious first reply to this objection is that CPMU should be statistically reformulated, roughly as follows: most individuals are mostly consistent. This reformulation would enable then a 'statistical intentional explanation' for now it is derivable that P(i,G) is probable.

A broader formulation of the objection excludes also this reply: the nomological or statistical premise is irrelevant, for, what does it matter for the intentional explanation of P(i,G) whether or not (almost) everybody is (almost) always consistent? The only consistency-question which can matter for this is whether or not i has been consistent in the particular case.

We take it that this objection has made it clear that there is something wrong with the NI-explication of our intuitions about intentional explanations. So let us make a new start by asking: What conditions seem to be intuitively necessary in order to allow us to say that P(i,G) is intentionally interpretable, i.e. can be explained intentionally, in terms of the pursuit of maximal utility, without making any assumption in advance concerning the precise structure of an intentional explanation? The following conditions seem plausible in any case: i wished maximal utility, i believed that participation was the best alternative in this respect, and i has been consistent. Note that the first two conditions imply that the third reduces to the fact that i participates; for, in the described circumstances he would have been inconsistent if and only if he had not participated.

The three mentioned conditions do not only seem necessary in order to justify the claim that P(i,G) is intentionally interpretable, but they also seem sufficient: it is at least difficult to think of other conditions we could or would like to add.

If this is so and if we continue to assume that an intentional explanation can be reconstrued as a deductive argument then it is plausible to assume that the three conditions function in this argument as the premises of the intentional explanation of P(i,G). Due to the already mentioned logical connection between the three, this set of premises reduces to WMU(i,G), DU-P(i,G) and P(i,G). Because this set includes P(i,G), it is evident that P(i,G) cannot be the conclusion of the underlying argument. For this would not only be trivial but it would also make the desiderative and epistemic premises redundant, and this is absurd.

Hence, our conclusion is that we have to look for a new explication of intentional explanations in which not only the desiderative and epistemic premises function truly as premises, but also the action-statement itself. In particular, we have to look for another, suitable conclusion.

In our (1983b) we have presented and defended extensively a general explication of everyday intentional explanation which satisfies these requirements. Here we will describe the general idea only briefly [7]. It will turn out that the present case of intentional-utilistic explanations forms a special case in two respects.

An intentional explanation of an action refers, by definition, to an intention, i.e. a conscious goal, of the actor. The crucial idea of the new explication is that the conclusion of the underlying argument of an intentional explanation is not that the actor has performed a certain action, but that the performed action was intentional, i.e. more or less consciously directed to a goal. In a complex premise, the explanation in the narrow sense, this goal is specified. Schematically, the argument reads as follows:

$$\frac{\text{i aimed at goal g with action a}}{\text{i performed a intentionally}}$$

The transition from the premise to the conclusion is, under any reasonable interpretation of the conclusion, an application of the rule of inference called Existential Generalization. The only thing we need to assume for this claim is that 'i performed a intentionally' can be explicitly defined as 'there is a goal which was aimed at by i with a'. In the present context we will speak of Intentional Generalization (IG) and of IG-explanations. The reverse transition, from the conclusion to the premise, is called Intentional Specification (IS), which is of course not a deductive but a synthetic step.

An IG-explanation will be called true or false (adequate or inadequate) if the premise is true or false, respectively.

It will be clear that it is possible that there is more than one true IG-explanation for one action.

In order to compare the present IG-explication with the current nomological (NI-)explication we will decompose the premise of an IG-explanation, i.e. i aimed at goal g with action a, into three components:

IG.1	i wished goal g
IG.2	i believed action a to be <u>useful</u> for g
IG.3	i performed a
	i performed a intentionally

With some temporal qualifications, which do not matter here, the conjunction of these components is equivalent to the original complex premise. The corresponding NI-explanation is the following:

NI.1	i wished goal g (i.e. the same as IG.1)
NI.2	i believed action a to be <u>necessary</u> for g
NI.3	some generalization of the conditional if NI.1 and NI.2 then i performs/performed a
	i performed a

It is easy to see that the three NI-premises together imply the three IG-premises, hence they imply the 'IG-conclusion'. Hence, in the light of the first implication there is a well-defined sense in which an NI-explanation is stronger than the corresponding IG-explanation.

This greater strength, however, is precisely responsible for the 'explicative objections' that have been raised against NI-explanations, especially by adherents of the so-called 'Practical Syllogism':

- the nomological premise (NI.3) is irrelevant,
- we do not claim by an everyday intentional explanation that the action could have been predicted,
- that the actor has performed a certain action is a presupposition when we start to look for an intentional explanation,
- frequently, alternative actions will have been possible, in which case the epistemic premise (NI.2) is much too strong.

All these objections do not apply to IG-explanations, as is easy to verify. Here we will only remark briefly a point which is of equal importance. The IG-explication rests on an unproblematic logical rule of inference. In this respect it is also highly superior to the Practical-Syllogism-explication (see note 6), which needs to assume a *magical* meaning postulate.

A straightforward example of a (non-utilistic) IG-explanation can be given for *participation*:

$U_X(i) > o$	i wishes the collective good X for G
$b_i(n) > o$	according to i, participation (in the pursuit of X) is useful for the production of X
P(i,G)	i participates in the pursuit of X for G
(Conj.)	i aims, with his participation, at X for G
(Int.Gen.)	i participates intentionally

It is of course impossible to construe an analogous IG-explanation for non-participation (unless we assume that the individual believes that non-participation is useful). However, it is possible to give intentional explanations for both actions, on a different level, viz. in terms of the goal 'maximal utility'.

The IG-explanation of P(i,G) in terms of maximal utility obviously is:

WMU(i,G)	i wishes maximal utility
DU-P(i,G)	i believes that participation is necessary and sufficient (hence useful) for maximal utility
P(i,G)	i participates in the pursuit of X for G
(Conj.)	i aims, with his participation, at maximal utility
(Int.Gen.)	i participates intentionally

Note that the premises are exactly those which we had derived earlier as the presumable premises. It is clear that an analogous explanation can be given for N(i,G). This type of explanation will be called *intentional-utilistic* (IU-)*explanation*.

IU-explanations are in two respects special cases of IG-explanations. In the first place, all IU-explanations refer to the same goal, viz. maximal utility. In the second place, the epistemic premise specifies that the relevant action is not only useful, but even necessary and sufficient.

We will call an action *IU-interpretable* if all premises of the corresponding IU-explanation are true. This makes it easy to formulate the answer to our original question *what* an intentional interpretation of maximization behaviour would look like, viz. the actions of the individuals are IU-interpretable. At the same time we get the answer to the question *when* an intentional interpretation of a utilistic reduction is realistic, viz. when the actions of the individuals are IU-interpretable, in which case we will speak of intentional-utilistic (IU-)reduction.

Our analysis should of course be such that the maximization hypothesis (MH) follows from the hypothesis that the actions of the individuals are IU-interpretable and this provides a nice touchstone for the analysis. The suggested intentional-utilistic interpretability hypothesis comes down to the following:

IUIH For all G and all i in G,
if P(i,G) then WMU(i,G) and DU-P(i,G), and
if N(i,G) then WMU(i,G) and DU-N(i,G)

It is easy to check that the following theorem is provable,

Th.5 IUIH iff WMU and MH

where WMU indicates that everybody wishes maximal utility in the P/N-choice, i.e.

WMU For all G and all i in G: WMU(i,G)

Besides the fact that IUIH implies MH, as was to be expected from an adequate account, it is also clear from Th.5 what IUIH adds to MH, viz. WMU. In words, utility maximization behaviour is intentionally interpretable if the individuals (consciously) wish maximal utility.

It is easy to see that the presented intentional interpretation of the maximization hypothesis designed for the utilistic reduction of Olson's hypothesis can be generalized to any context where a similar maximization hypothesis is useful in the reduction of social regularities.

Moreover, the scheme for utilistic reduction at the end of Section 2 can of course be extended to a scheme for intentional-utilistic reduction by adding on top the (deductive) 'neutralizing step' (NS):

$$\text{NS}\ \frac{\text{intentional-utilistic interpretability hypothesis}}{\text{maximization hypothesis}}$$

As has been said already, an intentional interpretation of utilistic reduction will not be realistic in every context. In some contexts, however, it will be realistic, viz. in situations where people are more or less forced to choose consciously between alternative actions. If the maximization hypothesis leads to the explanation of regularities in such contexts, then an intentional interpretation seems unavoidable. But there are of course also many contexts where there is no pressure for a conscious choice at all. If utilistic reduction is nevertheless possible in such a case, intentional interpretation may be wrong. How else to explain the maximization hypothesis in that case, we leave as an open question.

With respect to Olson's hypothesis itself, it is clear that there will be collective goods for which an intentional interpretation is realistic as well as goods for which this is not the case. However this may be, in Section 2 we have seen that utilistic reduction of this hypothesis on the basis of the maximization hypothesis is perfectly possible. In this subsection we have seen that an intentional interpretation *can* be given, that fits nicely with our intuitions concerning intentional explanation.

3.3. Types of transformation rules

In the context of sociology, transformation rules relate by definition collective behaviour with social phenomena. They will be indispensable ingredients in any type of heterogeneous reduction in sociology (and economics), whether or not the reduction is based on some form of utility theory.

In this final section we will discuss some types of transformation rules, including possible ways of explaining them. The following diagram represents the types to be considered:

Types of transformation rules	
analytic	
empirical	causal
	social
	institutional

With the global division into analytic and empirical transformation rules we subscribe to Lindenberg (1977). A transformation rule is *analytic*, and hence does not require explanation, if it follows from a definitional link between social phenomena and collective behaviour. For example, the degree of (internal) democracy can be defined as the degree of participation in decision processes. If all transformation rules in a case of reduction are analytic, this reduction is by definition homogeneous. Note that, in the context of collective goods, it is difficult to think of goods which are linked in this analytic way with collective behaviour.

Heterogeneous reduction presupposes at least one empirical, i.e. non-analytic, transformation rule, of which we will distinguish three types, called causal, social and institutional rules. These distinctions are made along the lines of the role of intervening subjects, i.e. subjects acting 'between' collective behaviour and social phenomena.

In the case of causal transformation rules there are no intervening subjects at all: the link between collective behaviour and social phenomena is purely a matter of causal processes. A typical example is: a decreasing number of smokers leads to a decreasing number of cases of lung cancer. Of course, the explanation of causal transformation rules is outside the scope of sociology.

Social transformation rules are by definition links following from actions of intervening subjects, which need not only be real human beings but also other types of acting subjects, like firms, cities, etc.. In most cases the transformation rules for collective goods will be of this type.

Except for the extreme case of institutional transformation rules (see below), social transformation rules presuppose that the collective behaviour creates circumstances in which the intervening subjects have some freedom of action. If these subjects maximize their expected utility in this 'space of action', it is of course possible that the apparent social regularity between their circumstances and the social phenomenon is again utilistically reducible, in which case new transformation rules will turn up.

For most social transformation rules not all these assumptions will be realistic. But this does not exclude explanations of such rules which perfectly fit into the program of methodological individualism, viz. explanations assuming intentional, but not necessarily utilistic, actions of the intervening subjects. In our ((1983c) Section 3.5) we have sketched two examples of such explanations for the transformation rules concerning collective goods.

We define institutional transformation rules as extreme cases of social transformation rules, viz. rules where the behaviour of the intervening subjects has been institutionalized. A typical example is the rule, in parliamentary democracy, that a bill gets the status of law if and only if it gets a majority of votes in parliament. Though there are intervening subjects involved, e.g. the president's signature, their role is fixed. Lindenberg (1977) seems to have thought only of this type of empirical transformation rules.

In many cases it will be obvious what institutional transformation rules are involved. In principle, however, the assessment of institutional rules is a matter of empirical research, think e.g. of historical or anthropological research. Suppose that the existence of such a rule has been extablished. Now, the quest for an explanation of that rule is in a sense meaningless. Of course, it makes sense to ask 'external' questions, like why (and how) the rule has been brought about, why it is maintained and why it is abolished or replaced. But it does not make sense to look for a 'deeper' mechanism which produces the institutional rule.

For this reason, institutional rules are very similar to the transformation rules in the reduction of the ideal gas law. If carefully formulated, the latter are, like 'water is H_2O', examples of ontological indentities: universal relations which need empirical justification but no (causal) explanation 8). Institutional rules differ only from ontological identities in that they are no universal connections, but artificial connections. This difference not only provides the reason why the indicated external questions make sense, it leaves after all some room for a distinction between sociology and physics 9).

Department of Philosophy
Westersingel 19
9718 CA Groningen
The Netherlands

NOTES

1) For empirical evidence for this hypothesis, see Olson (1965) and (1982)
2) The terms 'homogeneous' and 'heterogeneous' reduction were introduced by Nagel ((1961),Ch.11)
3) In our (1983a) we have argued that (advanced) textbooks in physics, though giving a correct account of the aggregation step for the ideal gas law, make a mental error with respect to the wall-attraction in the aggregation step leading to the more realistic law of Van der Waal's. It is also shown that this error can be avoided, but only by a complete reform of the aggregation step.
4) Our (1982) is in many respects a critical evaluation of Nagel's (1961) account of this example in the light of Causey's (1977) ontological approach to reduction.
5) It is precisely this aspect which gives the whole analysis

a static character: the individuals do not need to bother about the behaviour of the others. By changing the assumptions such that the (subjective) marginal effect depends, as a rule, on the number of participants the analysis would achieve a dynamic character. Taylor's (1976) 'supergames' might be useful for such an analysis. Moreover, important qualifications of the regularities seem unavoidable.

6) This is the crucial difference between the nomological explication and the 'semantic' explication of Von Wright in terms of the so-called Practical Syllogism. In the latter account of CPMU is (implicitly) assumed to be a meaning postulate.

7) There are also some differences in presentation. In particular in our (1983b), intentional-utilistic explanations are presented as intentional explanations of choices between alternative actions, to distinguish them from 'direct' intentional explanations of the chosen action. Unfortunately, this, in our opinion very adequate, distinction leads to many complicated formulations, which are certainly not suitable for a summary.

8) See note 4. Causey's distinction (Causey (1977)) between causal connections and identities is entirely based on whether or not an explanation is required. Like causal connections, identities need empirical justification. This distinguishes identities from analytic connections which need neither explanation, nor justification.

9) I like to thank the Netherlands Institute of Advanced Study (NIAS) at Wassenaar for the possibility of this research. I also thank Wolfgang Balzer, Carl Doerbecker (†), Henk Flap, Bert Hamminga and Henk Zandvoort for their comments on the draft of my (1983c), which forms the basis of this paper.

REFERENCES

Causey, R.L.: 1977, Unity of Science, Synthese Library 109, Reidel, Dordrecht.

Kuipers, T.A.F.: 1982, 'The reduction of phenomenological to to kinetic thermostatics', Philosophy of Science, 49.1,' 107-119.

Kuipers, T.A.F.: 1983a, 'The paradigm of concretization: the law of Van der Waals', to appear in Poznan Studies in the Philosophy of the Sciences and the Humanities.

Kuipers, T.A.F.: 1983b, 'The logic of intentional explanation', to appear in: The Logic of Discourse and the Logic of

Scientific Discovery (Eds. J. Hintikka and F. Vandamme).
Kuipers, T.A.F.: 1983c, 'Utilistische reductie in de sociologie', manuscript, partly to appear in Mens en Maatschappij, under the title 'Olson, Lindenberg en reductie in de sociologie'.
Lindenberg, S.: 1977, 'Individuelle Effekte, kollektive Phänomene und das Problem der Transformation', H. Eichner, W. Habermehl (Hrg.), Probleme der Erklärung sozialen Verhaltens, Meisenheim, Hain, 64-84.
Lindenberg, S.: 1982, 'De onvolledigheid van algemene hypothesen: een pleidooi voor verklarende modellen', Mens en Maatschappij, 57.4, 373-391.
Nagel, E.: 1961, The Structure of Science, Routledge, London.
Olson, M.: 1965, The Logic of Collective Action, Harvard UP, Cambridge.
Olson, M.: 1982, The Rise and Decline of Nations, Yale UP, New Haven.
Taylor, M.: 1976, Anarchy and Cooperation, Wiley London.